Management Science

Giving Feedback to Clients

Teck-Hua Ho, Catherine Yeung

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Giving Feedback to Clients

Teck-Hua Ho

Haas School of Business, University of California, Berkeley, Berkeley, California 94720; and
National University of Singapore, Singapore 119077, hoteck@haas.berkeley.edu

Catherine Yeung

NUS Business School, National University of Singapore, Singapore 119245, cyeung@nus.edu.sg

We examine a prevalent form of client–agent interaction through a feedback-giving game. In this game, a client undertakes a nontrivial task and is compensated based on her task performance, which is only made known to her when the client–agent interaction ends. Meanwhile, her performance is disclosed to an agent, who must then give the client feedback on her performance. Upon receiving the feedback, the client reports her happiness level, which in turn determines the agent's payoff. In eight studies involving 928 subjects, we vary the way the agent's cash earnings depend on the client's reported happiness. When we make the agent's earnings proportional to the client's reported happiness, the agent inflates his feedback, and the client reports a higher level of happiness (than that reported in a control condition where the agent always provides honest feedback). We show that neither the agent nor the client behaves altruistically in their reporting. The client reports being happier because she overestimates her performance and mistakenly believes that the agent's feedback is genuine. The agent stops inflating his feedback when doing so no longer benefits him. Finally, our main findings are shown to be robust with respect to several factors, including making the agent's feedback consequential in affecting the client's payoff. In summary, we show that the agent behaves opportunistically, and the client's overconfidence of her own performance is what makes this strategy successful.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2013.1846.

Keywords: opportunism; altruism; overconfidence; feedback giving game; client–agent interaction; behavioral economics
History: Received September 20, 2012; accepted August 29, 2013, by Uri Gneezy, behavioral economics. Published online in *Articles in Advance* February 12, 2014.

1. Introduction

Giving and receiving feedback are central to social interaction. People regularly give and receive feedback. For example, as professors, we give feedback to our students about their class performance, and we receive feedback from them about our teaching. Giving and receiving feedback are also important forms of communication because they can have significant emotional effects on the recipients. Positive feedback can affirm the worth of the recipients and make them feel happy and competent. Negative feedback frequently produces negative emotions and can generate a sense of incompetence in the recipients. In this research, we develop a simple feedback-giving game to study how people provide feedback to make others happy in social interactions.

Specifically, we investigate client–agent interactions in which the agent is in a position to provide feedback to the client based on some private information he has about the client's performance.[1] Because the client cares about her task performance, the agent's feedback can significantly affect her happiness. The agent wants to make the client happy because the agent's cash earnings will increase with the client's happiness.[2] There are many real-world examples of this type of social interaction.[3] For example, consider a

---

[1] For consistency and brevity, we use "she" to refer to the client and "he" to refer to the agent.

[2] Andrade and Ho (2007, 2009) show that an agent whose payoff is dependent on his client's decision often believes he will receive better treatment from a happier client. Their finding explains why agents are motivated to manage their clients' happiness when opportunities arise. When such opportunities involve feedback giving, the feedback becomes a natural platform for the strategic management of happiness.

[3] Previous research has documented such interactions in a wide range of organizational settings (Gordon 1996). People not only strategically please others in high-status supervisory positions by providing pleasant feedback, they also do the same to those in low-status subordinate positions, especially when they depend on the latter's specialized expertise. A similar kind of phenomenon occurs in market settings. For example, sales people please their customers to get them to buy their products. In these examples, the agents interact with their clients either only once or infrequently so that the specific feedback-giving episode is like a one-shot game. Our research is particularly relevant to interactions of this type, where there is little room for reputation building.

subordinate who must give feedback to his boss about the boss's performance in an inaugural speech. The subordinate cares about the boss's happiness because the boss is empowered to determine the subordinate's payoff through performance evaluation. The subordinate can imagine that the boss will experience unpleasant feelings upon receiving negative feedback. One way to mitigate these negative emotions is to provide inflated feedback and compliment the boss. However, the boss is keenly aware of the subordinate's incentive to inflate feedback and may not believe the feedback. This reaction may be anticipated by the subordinate, making it unclear, ex ante, whether he will inflate his feedback. It is also unclear whether feedback inflation will produce higher payoff than honest reporting, given the boss's likely skepticism about the feedback's truthfulness. If feedback inflation does increase payoff, why is this so? Finally, to what extent is feedback inflation driven by a self-interested motive to make more money, as opposed to an altruistic motive to make others happier? We examine these issues in an experimental game in which the participants engage in client–agent interactions.

We develop a new *feedback-giving game* to capture the abovementioned client–agent interaction. We simultaneously examine the feedback the agent provides and the client's response to the agent's feedback, when it is common knowledge that the client's reported happiness can directly influence the agent's cash earnings. In this game, the client undertakes a nontrivial task (solving 10 challenging math problems) in which she earns $1 for each correct answer. The client's performance is privately made known to the agent. The agent must send a feedback message ("You have got —— correct answers.") to the client. Upon receiving the message and before knowing her true performance, the client must indicate her level of happiness on an 11-point scale that ranges from 0 (not happy) to 10 (very happy). The agent earns $0.50 for each incremental point reported on this scale, so that the higher the client's reported happiness, the more money the agent makes.

This game captures two fundamental features of client–agent interactions in feedback giving and receiving. First, it is clear to the agent that his payoff depends on the client's reported happiness. Therefore, there is a strategic motivation for the agent to manage the client's happiness through giving feedback. Second, being aware of the agent's potential opportunistic motive, the client understands that the agent's feedback may not be truthful. These two defining features give rise to a number of possible scenarios. The agent may or may not inflate his feedback and the degree of inflation will vary, depending on the anticipated reaction of the client. In turn, the client's

reported happiness will vary as a function of how happy she is upon receiving the feedback and how much she believes the feedback.[4] We pin down the agent's feedback-giving strategy and the psychological mechanism underlying the client's response to the agent's feedback in eight studies by systematically manipulating the agent's payoff scheme.

In Study 1, we run the experimental condition where the agent's cash earnings increase linearly with the client's reported happiness and the control condition where the agent is incentivized to provide honest feedback. This study generates two empirical regularities: (1) the agent inflates his feedback, and (2) the client reports a higher level of happiness when the agent's payoff is tied to the client's reported happiness. As a consequence, the agent earns a higher monetary payoff. Before we further examine these empirical regularities, we conduct a new study to gain deeper insight into the way the client rates her happiness in our experimental setting where the rating is tied directly to the agent's payoff. In Study 2, the clients in one condition report their happiness, from which the agents' payoffs are determined (as in Study 1). The clients in another condition report their happiness and determine the agents' payoffs on a separate basis, such that the reported happiness does not affect the agents' payoffs. We do not observe any difference in the client's reported happiness across the two experimental conditions. Therefore, we conclude that the client reports the same level of happiness regardless of whether or not the agent's payoff is tied to the reported happiness.[5]

There are at least two competing explanations for the two empirical regularities identified in Studies 1 and 2. First, the *altruism hypothesis* posits that the agent inflates his feedback and tells a white lie to make the client happy and that the client reports a higher level of happiness to make the agent more money (especially when doing so does not cost her materially). Second, the *opportunism-overconfidence hypothesis* states that the agent inflates his feedback to make more money. The client is overconfident in estimating her own performance and, as a consequence, mistakenly believes that the (inflated) feedback is genuine and reports a higher level of happiness. We investigate these two hypotheses using the same feedback-giving game but change the way the agent is compensated.

Study 3 tests the *altruism hypothesis* by first examining the client's altruism in two experimental conditions. In the first condition, the agent's cash earnings

---

[4] The client's reported happiness can also be affected by other factors, such as her altruistic desire to make the agent more money. These possibilities are systematically tested and ruled out in our studies.

[5] We thank an anonymous reviewer for suggesting Study 2.

increase linearly with the client's reported happiness. In the second condition, the agent's cash earnings are also proportional to the client's reported happiness, but this payment rule applies only if the agent provides truthful feedback. Thus, the client is equally able to change the agent's earnings through her reported happiness in both cases, but she will receive less favorable feedback in the latter condition. If the client is purely altruistic, she will report the same level of happiness in both experimental conditions (as she simply wants to make the agent money). However, if her reported happiness is driven by her feelings toward the agent's feedback, she will report a lower level of happiness in the second condition. Because we observe a lower level of reported happiness in the second condition, we conclude that the client's reported happiness reflects her genuine feeling toward the feedback, and her reporting is not driven by an altruistic motive to make the agent more money. Study 3 also tests whether the agent inflates his feedback because he wants to tell a white lie to make the client happy. In a third experimental condition, we pay the agent a fixed fee for providing feedback (i.e., the agent's payment is independent of both the client's reported happiness and whether or not he is truthful). If the agent is altruistic and wants to tell a white lie to make the client happy, he will continue to inflate his feedback in this experimental condition. We find that the agent does not inflate his feedback in this condition. Overall, Study 3 rules out the altruism hypothesis.

Study 4 further examines the hypothesis that the client reports a higher level of happiness to make the agent more money (i.e., be altruistic to the agent) only when she thinks she will make more money than the agent; that is, the client's altruism is conditional on the premise that she is ahead of the agent in cash earnings. We test this *conditional altruism* hypothesis by running a condition where the client is paid $0.50 for each correctly answered math question and the agent is paid $1 for each incremental increase in reported happiness. Under this revised payment scheme, the client is behind the agent in terms of cash earnings. Because the client reports the same higher level of happiness as before, we rule out the explanation that the client reports a higher level of happiness only if she is ahead of the agent.

Studies 5 and 6 test the *opportunism-overconfidence* hypothesis. Study 5 examines why the client feels genuinely happier, despite her keen awareness of the agent's incentive to make her happy. In this study, we measure the client's estimate of her own performance after she has taken the task, but before she receives feedback from the agent. We find that the client overestimates her performance. In addition, the agent's feedback and the client's own performance estimate

are remarkably close. These results explain why feedback inflation leads to higher reported happiness for the client and a higher payoff for the agent, that is, the overconfident client mistakenly believes that the agent's inflated feedback is truthful.

Study 6 provides direct evidence that the agent acts opportunistically when providing inflated feedback. Specifically, this study tests whether the agent stops inflating his feedback when the opportunity to take advantage of the overconfident client is eliminated. In the first five studies, the agent can benefit from the client's overconfidence because he is compensated based on the client's reported happiness *before* she finds out her true performance. In Study 6, we make the agent's cash earnings proportional to the client's reported happiness ex post (her reported happiness *after* she has found out her true performance). Here, the agent stops inflating his feedback. Hence, we conclude that the agent is indeed opportunistic in providing feedback.

Studies 7A and 7B show the robustness of our findings in two separate contexts. In Study 7A, we conduct the feedback-giving game in abstract terms by eliminating the happiness reporting context from the experiment.[6] In Study 7B, we examine a context in which the agent's feedback is binary (e.g., "bad" versus "good" performance) such that feedback inflation entails sending a message that has the opposite meaning. In Study 8, we examine a setting in which the agent's feedback is consequential in that it can influence the client's decision on how she wishes to be compensated (a fixed fee independent of task performance versus a fee proportional to task performance) in an identical task performed right after receiving the agent's feedback. We continue to observe feedback inflation in all these settings.[7]
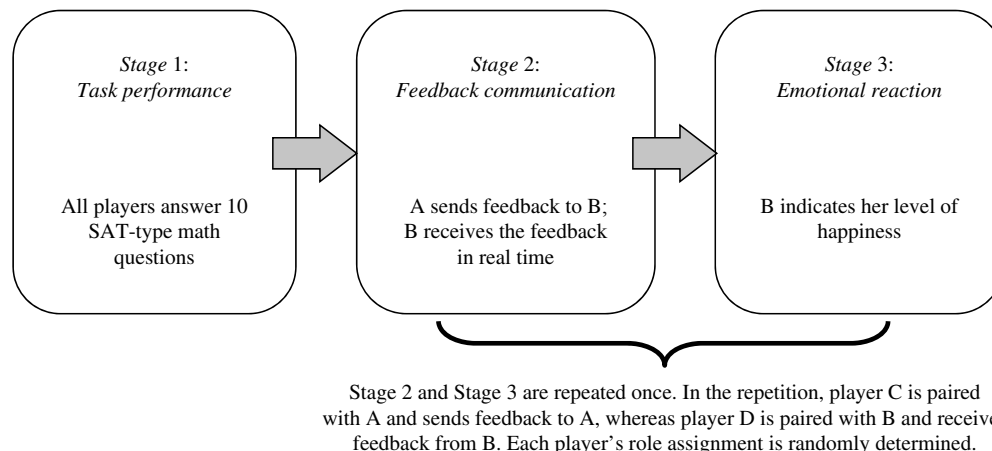
The remainder of this paper is organized as follows. Section 2 describes the feedback-giving game. Section 3 describes Studies 1 and 2 and summarizes the main empirical regularities. Section 4 develops the altruism and opportunism-overconfidence hypotheses and discusses how they can be tested in our experimental paradigm. Section 5 describes Studies 3 and 4, which serve to rule out the altruism hypothesis. Section 6 describes Studies 5 and 6, which test and support the opportunism-overconfidence hypothesis. Section 7 reports Studies 7A and 7B, which show that our finding on feedback inflation is robust with respect to two different contexts. Section 8 reports Study 8, which examines feedback giving when it is potentially consequential to the client's monetary

---

[6] See Camerer (2003) for a careful discussion of the benefits of studying people's strategic behavior in a generic (abstract) context.

[7] We thank the reviewers for making the excellent suggestion of conducting Studies 7A, 7B, and 8.

Figure 1   Three-Stage Experimental Paradigm for the Feedback-Giving Game

This figure illustrates the way four players (A, B, C, and D) participate in the feedback-giving game.
In an actual experiment, the number of players can be any multiples of four.

Stage 2 and Stage 3 are repeated once. In the repetition, player C is paired
with A and sends feedback to A, whereas player D is paired with B and receives
feedback from B. Each player's role assignment is randomly determined.

payoff. Finally, in §9, we discuss the implications of the results, highlight the limitations of our study, and suggest a number of future research directions.

2.   The Feedback-Giving Game

The feedback-giving game consists of three separate stages, namely, a task performance stage, a feedback communication stage, and an emotional reaction stage. In the task performance stage, each participant is asked to solve 10 SAT-type mathematics problems. The task is consequential because every correct answer earns the participant $1 so that he or she has the potential to earn up to $10 from the mathematics quiz. The participants only know their true performance and receive payment after the experiment is completed.[8]

In the feedback communication stage, the participants are randomly matched in pairs and assigned the role of either the agent or the client. The agent in the dyad first receives information about the client's actual performance on the mathematics quiz. The agent then sends feedback to the client anonymously and electronically. The possible feedback messages are "You have correctly answered [y] questions," where $y \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Let the client's actual number of correct answers be x. The agent's feedback is deflated if $x > y$, truthful if $x = y$, and inflated if $x < y$.

In the emotional reaction stage, the client responds to the feedback message by reporting her happiness on an 11-point scale ranging from 0 (not happy at all) to 10 (very happy). Note that the feedback message

sent by the agent is the only information the client receives before she reports her happiness. She does not know the actual number of correct answers, x, until the emotional reaction stage is completed. We repeat the feedback communication and emotional reaction stages, but match each participant with a different partner. Our experimental design ensures that each participant assumes the role of the agent and the client once and that the order of the role assignment is determined randomly. Figure 1 depicts the three-stage experimental paradigm.

Note that the same feedback-giving game paradigm is used in all eight studies reported in this paper. Table 1 provides an overview of these studies and their corresponding payment schemes for the agent. By simply employing different incentive schemes to compensate the agent or making simple modifications to the procedure, we are able to establish empirical regularities in feedback giving and receiving (Studies 1 and 2), test competing explanations (Studies 3–6), and extend our main findings to other agent–client interactions (Studies 7 and 8). Thus, our feedback-giving game provides a simple yet rich experimental paradigm for studying the social interaction between the agent and the client. For further discussion of why simple games are particularly useful for studying social interaction in an experimental setting, see Camerer (2003).

3.   Establishing the Empirical Regularities

3.1.   Study 1: Establishing Main Empirical Regularities

One hundred four undergraduates participated in this experiment. They were paid a show-up fee plus additional cash earnings contingent on their performance

---

[8] This is the case in all the studies except Study 6, where the clients are informed of their actual performance before they report their happiness.

Table 1   Experimental Objectives, Experimental Conditions, and Corresponding Payment Schemes Across the Eight Studies

Experiments and objectives	Experimental conditions and corresponding payment schemes		
Establishing the main empirical regularities and validating the happiness measure			
Study 1 　Establishing main empirical 　　regularities	*Proportional*: 　$0.50 × Client's reported happiness	*Truth-telling*: 　$0 if feedback is not truthful 　$3 if feedback is truthful	
Study 2 　Validating the combined 　　measure of client 　　happiness and agent payoff	*Proportional*: 　$0.50 × Client's reported happiness	*Truth-telling*: 　$0 if feedback is not truthful 　$3 if feedback is truthful	*Decoupled*: 　$0.50 × Client's reported 　　value on pay scale
Testing the altruism hypothesis			
Study 3 　Testing altruism	*Proportional*: 　$0.50 × Client's reported happiness	*Truth-telling, proportional*: 　$0 if feedback is not truthful, 　$0.50 × Client's reported 　　happiness if feedback is truthful	*Flat-fee*: 　$3 upon completion of 　　the feedback-giving task
Study 4 　Testing the client's 　　conditional altruism	*Proportional*: 　$0.50 × Client's reported happiness	*Swap incentive*: 　$1 × Client's reported happiness	
Testing the opportunism-overconfidence hypothesis			
Study 5 　Testing the client's 　　overconfidence	*Proportional*: 　$0.50 × Client's reported happiness	*Truth-telling*: 　$0 if feedback is not truthful 　$3 if feedback is truthful	
Study 6 　Testing the agent's 　　opportunism	*Proportional, ex post*: 　$0.50 × Client's ex post happiness 　　(i.e., reported after knowing 　　actual performance)	*Truth-telling*: 　$0 if feedback is not truthful 　$3 if feedback is truthful	
Feedback giving across contextual variations			
Study 7A 　Feedback-giving game 　　in abstract terms	*Proportional*: 　$0.50 × Client's reported happiness	*Abstract*: 　$0.50 × Client's reported value 　　on pay scale	
Study 7B 　Giving binary feedback 　　(good versus bad)	*Proportional*: 　$0.50 × Client's reported happiness	*Truth-telling*: 　$0 if feedback is not truthful 　$3 if feedback is truthful	
Giving consequential feedback			
Study 8 　Giving consequential feedback	*Proportional*: 　$0.50 × Client's reported happiness		

in the experiment.[9] The participants spent about one hour completing the tasks and earned $13 on average. They were randomly assigned to either one of the two experimental conditions: the proportional condition (in which the agent was paid based on the client's reported happiness) or the truth-telling condition (in which the agent was paid $3 whenever he reported the truth). For brevity, we refer to the former as the "prop condition" and the latter as the "truth condition." Both conditions followed an identical procedure except for the way the agent was paid. The experimental instructions for the prop condition are provided in the appendix.

3.1.1.   Procedure. The participants were randomly seated at cubicles equipped with computers. They were informed that they would be asked to complete a mathematics quiz and a feedback communication task, and that their payment would depend

on their actual performance on these two tasks. The participants were explicitly told that all communication during the experiment would be made through a Web-based software program, such that the participants could only interact electronically and anonymously.

The participants first completed the mathematics quiz. They were given 15 minutes to solve 10 SAT-type math problems. The difficulty of the problems was calibrated such that the average participant answered about five problems correctly. This design provided room for the agent to inflate or deflate the client's performance in the feedback communication. The participants were informed that every correct answer would earn them $1. Their answers were graded electronically by a software program. However, the participants were not able to access their actual scores until the experiment was over.

The participants were then randomly matched into pairs and engaged in feedback communication about each other's quiz performance. They were informed

---

[9] Cash payments were paid in Singapore dollars (SGD 1 = USD 0.8) in all studies reported in this paper.

that they would be asked to engage in two rounds of feedback communication and that they would be paired with a different participant in each round. In one round, they would act as the agent by providing feedback, and in the other round they would act as the client by receiving feedback and reporting their own level of happiness. The order of their role assignment was determined randomly, so that half the subjects acted as the agent first and the other half as the client first. Each agent then received a report showing his client's performance in the mathematics quiz. The agent was required to send feedback to his client by completing the statement "You have correctly answered ____ questions." Upon receiving the message, the client indicated her happiness on an 11-point scale ranging from 0 (not happy at all) to 10 (very happy). This ended the first round of feedback giving. In the second round, each participant was matched with a different partner and assumed a different role. When the second round ended, the participants were informed of their total payment from the mathematics quiz and the feedback communication task. The participants were paid in cash before leaving the experiment.

Payment Scheme. In the prop condition, the agents were paid based on their clients' reported happiness; each higher scale point yielded an additional $0.50. In the truth condition, the agents received $3 for truthful reporting and nothing ($0) for untruthful reporting.

To ensure that the participants fully understood the payment scheme and the experimental instructions, we conducted an understanding test and required the participants to answer all the questions correctly before they could proceed to the feedback-giving game. All of the participants passed the test.

3.1.2. Results and Discussion. *Manipulation and Confound Checks.* We pooled the data from the two rounds of feedback communication in our analysis of the agent's feedback[10] and the client's reported happiness.[11] The top panel of Table 2 summarizes the main results. The participants' actual quiz performance did

---

[10] We conducted the following analyses to examine whether the agent's feedback differed as a function of the order of role assignment (i.e., whether a participant played the agent role first or the client role first). For each study, we ran an analysis of variance (ANOVA) with the agent's feedback as the dependent variable and the order of role assignment and the experimental conditions as the independent variables. The *p*-values for the main effect of role assignment order ranged from 0.27 to 0.93 (average, 0.57) across all the studies. The *p*-values for the interaction effect ranged from 0.13 to 0.98 (average, 0.66). Hence, we pooled the data from the two rounds of feedback communication in the analysis of the agent's feedback.

[11] Similarly, for each experiment, we ran an ANOVA with the client's reported happiness level as the dependent variable and the order of role assignment and the experimental conditions as the independent variables. The *p*-values for the main effect of role

Figure 2   Scatter Plot of Agent's Feedback Against Client's Actual Quiz Performance in the Prop Condition in Study 1
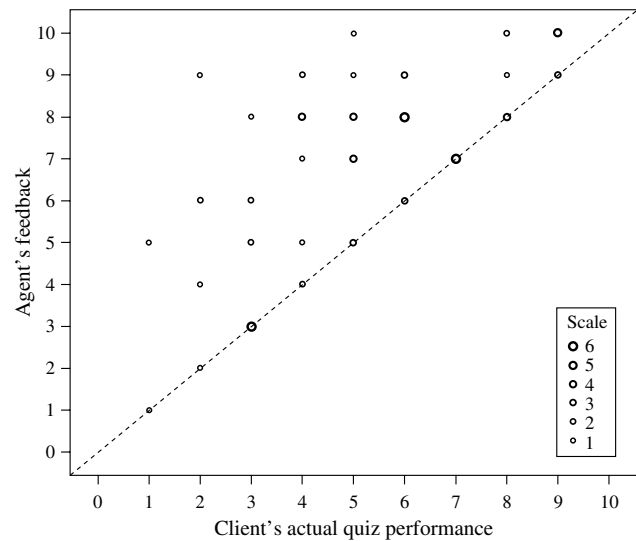
Note. The dotted line represents the 45° line: feedback = actual quiz performance.

not differ in the two experimental conditions ($M_{prop} = 5.25$ and $M_{truth} = 4.88$; $F(1, 102) < 1$). In the truth condition, all the agents reported truthfully. Therefore, we were able to use the truth condition as a no-inflation benchmark for comparison.

Feedback Inflation. In the prop condition, 59.4% of the agents inflated their feedback, and the remaining agents reported honestly. The average feedback in the prop condition was 6.97, which was 2.09 units higher than that in the truth condition ($M_{prop} = 6.97$ versus $M_{truth} = 4.88$; $F(1, 102) = 25.0$, $p < 0.001$). Figure 2 provides a scatter plot of the agent's feedback against the client's actual performance in the prop condition. If the agents had reported truthfully, all the data points would have fallen on the 45° dotted line. As shown, the majority of the points are above the 45° line, indicating that the majority of agents inflated their feedback.

Client's Reported Happiness. The clients in the prop condition reported a higher level of happiness than their counterparts in the truth condition ($M_{prop} = 6.25$ versus $M_{truth} = 4.18$; $F(1, 102) = 11.60$; $p = 0.001$). To examine whether this difference was driven primarily by the difference in feedback across the two

---

assignment order ranged from 0.25 to 0.99 (average, 0.71) across all of the studies. The *p*-values for the interaction ranged from 0.13 to 0.93 (average, 0.54) in all studies except for Study 2, where a marginally significant interaction effect was observed ($p = 0.06$). To check whether the order of role assignment affected the main results in Study 2, we also conducted statistical tests with role assignment order added to the statistical models. These additional analyses yielded the same results as those reported in this paper. For the sake of consistency and brevity, we pooled the data from the two rounds of feedback communication in all of our analyses of the client's reported happiness.

Table 2   Summary of Results (Study 1–Study 8)

	Math quiz performance	Agent's feedback	Client's reported happiness	
Study 1 (N = 104)				
Proportional (n = 64)	5.25 (2.22)	6.97 (2.17)	6.25 (3.25)	
Truth-telling (n = 40)	4.88 (1.92)	4.88 (1.92)	4.18 (2.61)	
Study 2 (N = 136)				
Proportional (n = 44)	4.84 (1.94)	6.00 (2.04)	6.14 (2.65)	
Truth-telling (n = 40)	5.23 (1.99)	5.27 (1.83)	4.63 (2.76)	
Decoupled (n = 52)	5.15 (2.35)	6.35 (2.41)	6.35 (2.99)	
Study 3 (N = 140)				
Proportional (n = 60)	4.92 (1.72)	6.47 (2.17)	6.33 (3.26)	
Flat-fee (n = 24)	4.75 (1.75)	4.83 (1.81)	4.37 (2.67)	
Truth-telling, proportional (n = 56)	4.93 (1.69)	4.93 (1.69)	4.91 (2.97)	
Study 4 (N = 92)				
Proportional (n = 48)	4.33 (2.14)	6.02 (2.38)	7.08 (3.07)	
Swap payoff (n = 44)	5.05 (2.21)	6.27 (2.46)	6.43 (3.44)	
Study 5 (N = 80)				
Proportional (n = 40)	4.93[a] (2.45)	6.08 (2.43)	6.28 (3.27)	
Truth-telling (n = 40)	5.07 (1.98)	5.07 (1.98)	4.25 (2.86)	
Study 6 (N = 96)			First	Second
Proportional, ex post (n = 64)	5.66 (2.09)	5.39 (2.46)	5.09 (3.17)	5.73 (3.15)
Truth-telling (n = 32)	5.03 (2.51)	5.03 (2.51)	4.34 (3.07)	4.22 (3.25)
Study 7A (N = 84)				
Proportional (n = 40)	4.68 (1.86)	6.43 (2.22)	5.43 (2.86)	
Abstract (n = 44)	4.30 (1.95)	5.23 (2.61)	7.25 (2.96)[b]	

Study 7B (N = 116)

	Math quiz performance		Client's reported happiness
Proportional (n = 60)	5.33 (1.96)		7.22 (3.34)

		Feedback	
		Bad	Good
Actual			
	Bad	17	18
	Good	2	23

	Math quiz performance		Client's reported happiness
Truth-telling (n = 56)	5.79 (2.16)		6.14 (3.15)

		Feedback	
		Bad	Good
Actual			
	Bad	23	0
	Good	0	33

	Math quiz performance	Agent's feedback	Client's reported happiness
Study 8 (N = 80)			
Proportional (n = 80)	Quiz 1: 5.51 (2.17)	6.69 (2.49)	6.64 (3.56)
	Quiz 2: 6.81 (2.56)		

Note. Numbers in parentheses denote standard deviations.
[a]The client's prediction of her own performance was 6.23 (SD = 2.53).
[b]The number represents the client's decision on the agent's payment, as no happiness label was used.

experimental conditions, we ran an analysis of covariance (ANCOVA) with the agent's feedback added as a covariate. As expected, the effect of feedback was highly significant ($F(1, 101) = 10.60$; $p = 0.002$), whereas that of the payoff scheme manipulation became only marginally significant ($F(1, 101) = 3.06$; $p = 0.083$).[12] This suggests that the clients in the prop condition reported higher levels of happiness, and

that this increase was largely attributed to the higher (inflated) feedback they received from the agents. Note that the higher reported happiness resulted in an average payoff of $3.13. This result implies that feedback inflation pays, because if the agent had reported

---

[12] We also conducted the same analysis for all subsequent studies. In each study, the inclusion of feedback as a covariate dramatically reduced the statistical significance of payoff scheme manipulation,

and the effect of feedback on reported happiness was always highly significant (always less than 0.005). These results suggest that payoff scheme manipulation influences the client's reported happiness mainly through modifying the agent's feedback. For the sake of brevity, we do not report this analysis repeatedly in this paper. The detailed results are available from the authors upon request.

truthfully, then the client would have indicated a lower happiness rating of 4.18 (as in the truth condition). This would have resulted in an average payoff of $2.09, 33% lower than the actual payoff of $3.13.

Study 1 generated two empirical regularities: (1) the agent inflates his feedback, and (2) the client reports a higher level of happiness when the agent's payoff is tied to the client's reported happiness. Before we proceed to examine the mechanisms underlying these empirical regularities, we must first verify whether tying the agent's payoff to the client's reported happiness in a *combined* measure changes the way the client reports her happiness. Specifically, will the client report her happiness the same way when her reported happiness bears no consequence for the agent's payoff? The answer to this question would facilitate a more precise interpretation of subsequent findings on the client's reported happiness.

3.2. Study 2: Validating the Combined Measure of Client Happiness and Agent Payoff

One hundred thirty-six participants were randomly assigned to one of the following three conditions: the prop condition, the truth condition, or a new "decoupled condition." The decoupled condition was the same as the prop condition, except that the client was asked to report her happiness and her decision on the agent's payoff on two separate scales. Specifically, the client first reported her happiness on an 11-point scale ($0 =$ not happy at all; $10 =$ very happy), and then indicated her decision on how much the agent should be compensated on a separate scale ranging from 0 to 10, knowing that each scale point would earn the agent $0.50. Because the client's reporting of happiness is decoupled from her decision on the agent's payoff, her reporting should be free from consideration of the agent's payoff. Comparison of this "decoupled" reported happiness and that in the prop condition should reveal whether the tying of the agent's payoff to the client's reported happiness leads to any distortion in the client's reported happiness.

3.2.1. Results. The second panel of Table 2 shows the experimental results. Again, the math quiz performance did not differ across the three conditions ($F < 1$). Before comparing the client's reported happiness in the decoupled and prop conditions, we first had to ensure that the clients received the same level of feedback from their agents in the decoupled and the prop conditions and that the average feedback in the two conditions was higher than that in the truth condition. Indeed, the agents' feedback in the prop and decoupled conditions was statistically not different ($M_{prop} = 6.00$ and $M_{decoupled} = 6.35$; $F < 1$; $p = 0.43$). Moreover, the average feedback in the two conditions was statistically higher than that in the truth condition (6.18 versus 5.27; $F(1, 133) = 5.01$; $p < 0.05$).

The clients' reported happiness did not differ between the decoupled and the prop conditions ($M_{prop} = 6.14$ and $M_{decoupled} = 6.35$; $F < 1$; $p > 0.50$), and the clients in these two conditions reported being happier than their counterparts in the truth condition (6.25 versus 4.63; $F(1, 133) = 9.29$; $p < 0.05$). In combination, these findings suggest that tying the agent's payoff to the client's reported happiness does not change the way the client reports her happiness.

When the clients in the decoupled condition were asked to determine how much their agents should be compensated, the average rating was 7.73, earning the agents a higher payoff than those in the prop condition ($M_{decoupled} = \$3.87$ versus $M_{prop} = \$3.07$; $F(1, 94) = 7.11$; $p < 0.01$).[13] Note that the clients in the prop condition could have made the agents as much money as those in the decoupled condition did, simply by inflating their reported happiness. However, they did not do so. Hence, there appears to be a preference among the clients to report their happiness truthfully.

3.2.2. Discussion. One novel aspect of the feedback-giving game is the tying of the agent's payoff to the client's reported happiness. This design feature plays an important role by making it clear to the agent that his payoff is determined by the client's reported happiness, and hence that it is in his best interest to manage the client's happiness through the provision of carefully considered feedback. The same feature also makes it clear to the client that the agent's feedback may be strategic and inflated. Importantly, Study 2 shows that although this coupling might prompt such strategic considerations, it did not distort the client's reported happiness; that is, the level of reported happiness was the same regardless of whether the agent's payoff and the client's reported happiness were coupled or decoupled.

In summary, the results thus far suggest that when the agent's payoff is proportional to the client's reported happiness, the agent inflates his feedback. This feedback inflation results in a happier client and higher earnings for the agent. This finding raises two fundamental questions. First, why did the agent inflate his feedback? Second, why did the client respond favorably to the feedback even though she was keenly aware that the agent had incentive to inflate? To answer these questions, we review the existing research that bears relevance to our feedback-giving game. We then develop alternative hypotheses, which we test in Studies 3–6.

---

[13] The higher average payment was driven by half the clients who chose the maximum scale point when they determined the agent's payoff. The other half simply chose a rating close to their reported happiness.

4. Understanding Feedback Giving and Receiving

There are two possible theories underlying the empirical regularities established in Studies 1 and 2. The first theory assumes that people are *opportunistic* and their behavior is guided primarily by self-interest. As a consequence, these people choose actions that serve to maximize their own payoffs. The second theory assumes that people are *altruistic* and care about the welfare of others. They derive utility when the other's welfare improves and hence are willing to choose actions that increase this welfare (Andreoni 1990, Sugden 1982). In the following sections, we elaborate on how these two theories can account for the main empirical regularities. Note that these theories can lead to distinctly different predictions when different payment schemes are used to compensate the agent. Therefore, we systematically vary the payment scheme in our subsequent experiments to separate the two theories and determine which theory is better at explaining the main empirical regularities.

4.1. Opportunism-Overconfidence Hypothesis

The agent may have acted opportunistically in inflating his feedback. If the agent anticipates that the client will be happier when she thinks she performed well in the quiz, it is in the agent's self-interest to enhance the feedback to please the client and earn more money. However, the client is keenly aware of the agent's incentive to please her. Therefore, even if she receives favorable feedback, she may not necessarily report a higher level of happiness because she may not believe the feedback. Thus, whether or not the client believes the favorable feedback determines whether the agent's feedback inflation will work.

Research in psychology suggests that people generally desire favorable feedback because such feedback enhances their self-conception and makes them feel good about themselves (Sedikides 1993, Taylor and Brown 1988). As a result, people tend to overweight positive relative to negative feedback, even in situations where they are incentivized to judge themselves accurately (Eil and Rao 2011, Mobius et al. 2011). A positive self-conception also leads people to be overconfident about their own performance in challenging tasks (Moore and Healy 2008). In our experiment, an overconfident client may believe she answered five questions correctly when in fact only four were answered correctly.[14] When

an overconfident client receives favorable feedback, she may believe that the feedback is truthful, despite knowing that the agent may be opportunistic. Accordingly, we posit that if the degree of feedback inflation closely matches with the level of overestimation, the client will believe the feedback and report a higher level of happiness.

Together, the agent's opportunism and the client's overconfidence can explain why the agent inflates his feedback and why inflation works to increase the client's reported happiness. Specifically, the agent opportunistically inflates his feedback to increase the client's reported happiness and earn more money for himself. This strategy works because the client is overconfident and mistakenly believes the feedback is genuine, and hence reports a higher level of happiness. We refer to this theory as the *opportunism-overconfidence* hypothesis.

4.2. Altruism Hypothesis

The agent and the client may have acted altruistically in the feedback-giving game. The agent may have inflated his feedback because he wanted to make the client happy, and, similarly, the client may have reported a higher level of happiness because she wanted the agent to make more money. The extensive literature on dictator games provides compelling evidence in support of this altruism hypothesis. In the standard dictator game, a dictator determines the allocation of a fixed pot of money between herself and another person (an anonymous responder). Because the responder cannot reject the dictator's decision, the dictator has absolute power to divide the pot of money in any way she pleases. If the dictator is purely self-interested, she will maximize her own payoff by making the smallest possible offer to the responder (i.e., \$0). However, the dictator frequently gives about 20% of the pot to the responder, which in effect reduces her own payoff by 20% (Forsythe et al. 1994). Indeed, numerous dictator game experiments report positive dictator offers (for a comprehensive review, see Camerer 2003), which suggests that people care about others' payoff and are willing to give others money even if it reduces their own payoff.

We reason that if altruism occurs when people have to bear a direct material cost (as in the dictator game), it should occur more readily when people do not have to bear any material cost to be altruistic. This is potentially the case for the client in the feedback-giving game. When the client knows that she can earn the agent more money by simply reporting a higher level of happiness, at no cost to herself, she may do so to benefit the agent.

It is also possible that the client cares about the agent's payoff *only* when she is ahead of the agent (as the case of the dictator). Specifically, if the client

---

[14] The overestimation of one's own performance has been distinguished from other forms of overconfidence, such as overplacement (overplacing one's own performance relative to that of others due to the belief that one is better than others) and overprecision (one is overly certain of one's own performance; Moore and Healy 2008). People can also exhibit other types of error and bias in estimating their own performance (see Kruger and Burrus 2004, Moore and Small 2007, Moore et al. 1999).

thinks that she is going to make more money than the agent (which is likely in our feedback-giving game), she may choose to be altruistic by overstating her happiness to benefit the agent.[15] We refer to this altruistic preference as *conditional altruism*, because the client's altruism is conditional on the premise that she is ahead of the agent (i.e., she makes more money than the agent). This conditional altruism is different from the *unconditional altruism* discussed above in that the latter form of altruism can occur without the client having an advantageous payoff position over the agent.

The agent's feedback inflation may also be driven by an altruistic motive. If the agent anticipates that telling the client her true performance may cause her to be unhappy, he may choose to sugarcoat the feedback. However, this inflated feedback could be costly because the client may wrongly interpret the agent's good intention as a self-interested strategy to make more money. As a consequence, the agent's altruism may backfire and cause him to receive a lower payoff from a suspicious client. Nevertheless, research shows that people are sometimes willing to tell an altruistic white lie to improve the welfare of another, even if their white lie slightly reduces their own payoff (Erat and Gneezy 2012; see also Gneezy 2005). Moreover, the abovementioned robust finding on dictator games suggests that people are willing to make a sacrifice if they believe their action will benefit others. Hence, the agent's decision to inflate his feedback may be driven by an altruistic motive to make the client happier, even if this altruistic act is costly to implement.

The altruism and the opportunism-overconfidence hypotheses both offer compelling explanations for the agent's decision to inflate his feedback and the client's favorable response to such feedback. In Studies 3 to 6, we systematically vary the agent's payment scheme to determine which hypothesis best explains the observed empirical regularities.

5. Testing the Altruism Hypothesis

5.1. Study 3: Testing Altruism

Study 3 consisted of three experimental conditions. We included the usual prop condition like before. To test whether the agent inflated his feedback simply to make the client happier, we ran a new experimental condition in which the agent was paid a flat fee of $3. In this flat-fee condition, the agent could choose to give honest feedback, inflated feedback, or deflated feedback, without having to worry about the payoff consequences. Put differently, the agent had complete freedom to influence the client's happiness through

the feedback he provided, while keeping his payoff constant. If the feedback inflation is driven by an altruistic motive to make the client happier, the agent should be at least as, if not more, likely to inflate his feedback in this "flat-fee" condition as in the prop condition.

We also studied whether the client reported a higher level of happiness because she wanted to make the agent more money. To do so, we ran a "truth-telling, proportional" condition ("truth-p condition" hereafter), in which the agent was paid based on the client's reported happiness, but received the payoff only if he reported truthfully. This design eliminated the agent's incentive to inflate his feedback, while allowing the client to be altruistic by reporting a higher level of happiness and hence making the agent more money. If the client's reported happiness was mainly driven by the agent's feedback, the level of reported happiness would decrease in the truth-p condition with less favorable (yet truthful) feedback. We would not expect such a decrease in reported happiness if the reporting was driven by an altruistic motive to make the agent more money. In fact, in the latter case, one could even argue that higher reported happiness could arise in the truth-p condition because the client would be quite certain that the agent was reporting truthfully.

One hundred forty participants were randomly assigned to one of the three conditions: (1) the prop condition, (2) the flat-fee condition, and (3) the truth-p condition. The procedure was similar to that employed in Study 1 except that the agents in the two latter conditions were told differently about the ways they would be paid.

5.1.1. Results. The findings from Study 3 are shown in Table 2. As expected, the math quiz performance did not differ across the three conditions ($F < 1$). Next, we report two separate sets of statistical comparisons that test the altruism hypothesis. We first compare the prop and the flat-fee conditions to examine the agent's altruistic behavior, and then we compare the prop and the truth-p conditions to investigate the client's altruistic behavior.

Are Agents Altruistic in Giving Feedback? In the prop condition, 48.3% of the agents inflated their feedback, and the remaining 51.7% reported truthfully. In the flat-fee condition, 83.3% of the agents reported truthfully, 8.3% of the agents inflated their feedback, and the remaining 8.3% of the agents deflated their feedback. The feedback was less favorable in the flat-fee condition than in the prop condition ($M_{prop} = 6.47$ versus $M_{flat-fee} = 4.83$; $F(1, 82) = 10.65$; $p < 0.001$). In fact, the feedback provided in the flat-fee condition was not statistically different from the client's actual quiz performance ($M_{feedback} = 4.83$; $M_{quiz} = 4.75$; $t(23) = 0.62$; $p > 0.50$). Thus, the agents stopped inflating their

---

[15] See Bolton and Ockenfels (2000) for a review of related works on equity and distribution preferences.

feedback when their monetary payoff was not tied to the clients' reported happiness. Consequently, we conclude that the agent's decision to inflate his feedback was not driven by an altruistic motive to make the client happy.

Are Clients Altruistic in Responding to Agents' Feedback? As expected, no agent inflated in the truth-p condition. The average feedback in the truth-p condition ($M = 4.93$) was statistically lower than that in the prop condition ($M = 6.47$; $F(1, 114) = 17.97$; $p < 0.001$). Correspondingly, the client reported a lower level of happiness in the truth-p condition than in the prop condition (4.91 versus 6.33; $F(1, 114) = 6.01$; $p = 0.01$). If the client's happiness reporting was driven by her desire to make the agent more money, the level of reported happiness would not have decreased in the truth-p condition. However, this was not the case. The client reported a lower level of happiness when the feedback she received was less favorable (yet truthful), despite her knowing that a lower level of reported happiness would mean a lower payoff for the agent. Hence, we rule out the explanation that the client has a desire to improve the agent's payoff.

5.2. Study 4: Testing the Client's Conditional Altruism

Study 4 examined another form of altruism that may have been displayed by the client. The client may have an altruistic desire to make the agent more money (by overstating her happiness rating) only if she thinks she is going to make more money than the agent. We refer to this altruistic preference as *conditional altruism*, because the client's altruism is conditional on the premise that she is ahead of the agent in terms of payoff. In the feedback-giving games we have reported thus far, conditional altruism could have occurred on the client's side because the payoff scheme was structured in such a way that the client was likely to make more money than the agent. Thus, in Study 4, we tested the conditional altruism hypothesis by reversing the payoff scheme of the client and the agent and then examining how this swap manipulation changed the client's reported happiness. Specifically, we ran the prop condition as before and a new "swap payoff" condition, in which the payoff schemes for the client and the agent were reversed (that is, the client was paid only $0.50 for each correctly answered math question, but the agent was paid $1 for each incremental increase in happiness reported by the client). Hence, in the new "swap payoff" condition, the agent was likely to be ahead of the client in terms of payoff. If the client indeed cares about how her payoff compares to the agent's, and if her happiness reporting is driven by this consideration, as the conditional altruism hypothesis suggests,

we would observe higher reported happiness in the prop condition than in the swap payoff condition.[16]

Ninety-six participants were randomly assigned to one of the two experimental conditions.[17] The experimental procedure was identical in the two conditions (as in the previous experiments), except that the client in the swap payoff condition was paid only $0.50 for each correctly answered question, and the agent was paid $1 for each point of happiness reported by the client.

5.2.1. Results. The findings from Study 4 are presented in Table 2. Neither the client's math quiz performance ($M_{prop} = 4.33$ and $M_{swap} = 5.05$; $F(1, 90) = 2.47$; $p = 0.12$) nor the agent's feedback on her performance ($M_{prop} = 6.02$ and $M_{swap} = 6.27$; $F < 1$; $p > 0.50$) differed between the two experimental conditions. More importantly, the client's reported happiness did not differ between the two conditions ($M_{prop} = 7.08$ and $M_{swap} = 6.43$; $F(1, 90) = 0.92$; $p = 0.34$). To ensure that the client's reported happiness was still driven mainly by the agent's feedback, we ran an ANCOVA with the agent's feedback added as a covariate. As expected, the effect of feedback on reported happiness was highly significant ($F(1, 89) = 10.67$; $p = 0.002$). In combination, these results suggest that the client's reported happiness was driven mainly by the favorability of the agent's feedback, but not by a comparison of her and the agent's payoff. Hence, the conditional altruism hypothesis was not supported.

Thus far, our findings suggest that when the agent's payoff is tied to the client's reported happiness, the agent inflates his feedback. This inflation results in higher happiness reported by the client, which in turn increases the agent's payoff. Studies 3 and 4 specifically show that the agent does not tell a white lie to make the client happier, and the client does not deliberately report a higher happiness rating to make the agent more money. These findings are consistent with those of Study 2, which suggest that the client reports her happiness as if she pays no attention to how her reported happiness may affect the agent's payoff (i.e., she reports the same level of happiness as her counterpart in the decoupled condition). These findings, in combination, provide strong evidence that the client's happiness reporting is not influenced by the altruistic motive of making the agent more money.

---

[16] In fact, the client might even deliberately report an exceptionally low happiness rating in the swap payoff condition because by doing so she could make the agent less money and reduce the disparity between her and the agent's monetary payoffs.

[17] One participant did not provide any feedback when playing the agent role, and the corresponding client reported happiness without having received any feedback. The data from these two participants and the other two who were matched with them were removed from data analysis, resulting in 92 usable data points. This problem did not occur in any other studies reported in this paper.

6. Testing the Opportunism-Overconfidence Hypothesis

Studies 5 and 6 test the opportunism-overconfidence hypothesis. In Study 5, we explore why the client reports a higher level of happiness, despite being keenly aware of the agent's incentive to inflate his feedback. In Study 6, we test the hypothesis that the agent is opportunistic and inflates his feedback to make more money.

6.1. Study 5: Testing the Client's Overconfidence

The objective of Study 5 is to provide insights into why the client reports a higher level of happiness. The opportunism-overconfidence hypothesis posits that the client believes the agent's (inflated) feedback and reports a higher level of happiness because she overestimates her own performance, and this overestimation is roughly the same as the degree of feedback inflation.

Study 5 consisted of two experimental conditions: (1) the prop condition and (2) the truth-telling condition (i.e., the same set of conditions as in Study 1). The experimental procedure and payment schemes for these conditions were identical to those employed in Study 1, except the participants in the prop condition were asked to estimate their actual performance immediately after they had taken the quiz. The prediction task was consequential in that an accurate prediction would earn the participants a bonus payment of $0.50. Hence, the participants were motivated financially to estimate accurately.

6.1.1. Results. In the prop condition, 40% of the agents inflated their feedback, and the remaining 60% reported truthfully. Also, as expected, no feedback inflation was observed in the truth condition (i.e., all agents reported truthfully). The average feedback in the prop condition ($M = 6.08$) was higher than that in the truth condition ($M = 5.07$; $F(1, 78) = 4.07$; $p = 0.05$). The client's reported happiness was higher in the prop condition ($M_{\text{prop}} = 6.28$ versus $M_{\text{truth}} = 4.25$; $F(1, 78) = 8.68$; $p = 0.004$), resulting in a higher payoff for the agent. These differences cannot be attributed to task performance because the participants in the two conditions did not differ in their quiz performance ($M_{\text{truth}} = 5.07$ and $M_{\text{prop}} = 4.93$; $F(1, 78) < 1$).

Performance Estimation and Feedback Inflation. Consistent with the opportunism-overconfidence hypothesis, the clients overestimated their math quiz performance. Their estimated number of correct answers was 6.23, which was statistically higher than their actual number of correct answers ($M = 4.93$; $t(39) = 5.18$; $p < 0.001$).

We next assessed how closely the agent's (inflated) feedback matched with the client's estimate of her own performance. At the aggregate level, the agent's average feedback ($M = 6.08$) was remarkably close to the client's average performance estimate ($M = 6.23$). To examine whether the agent's feedback matched the client's estimate at the individual level, we computed the difference between the two values for each client–agent interaction. The average difference was 0.15, which was statistically not different from zero ($t(39) = 0.42$; $p > 0.50$). In addition, we fitted a regression line using the client's estimate as the x-variable and the agent's feedback as the y-variable. The best fitted line that passes through the origin (represented by the solid line in Figure 3) has a slope of 0.92 ($t(39) = 17.84$; $p < 0.001$; $R^2 = 0.89$).[18] Figure 3 shows a scatter plot of these two variables with the solid line representing the best fitted line and the dotted line representing the 45° line. Overall, these findings indicate that the agent's feedback almost perfectly matches the client's performance estimate, which in turn explains why the feedback is considered credible by the client.

The opportunism-overconfidence hypothesis suggests that an opportunistic agent will inflate his feedback only if he thinks that he can benefit by doing so (i.e., the client reports her happiness *before* finding out her actual performance). Would the opportunistic agent inflate his feedback if he knew that the client reports her happiness *after* she finds out her true performance (i.e., *ex post* happiness)? Clearly, the agent will stop inflating his feedback if he believes feedback inflation will make the client unhappy and report a lower level of ex post happiness. The agent should truthfully report if he believes the client will prefer a truthful message. Alternatively, the agent may wish to underreport if he believes that a pleasant surprise will help to increase the client's ex post happiness. Hence, a direct way to test whether the agent is indeed opportunistic is to check whether the agent will stop inflating his feedback if his earnings are proportional to the client's ex post happiness. In Study 6, we provide a further test of the opportunism-overconfidence

---

[18] We also examined the client's reaction when the agent inflated his feedback *beyond* the client's performance estimate. We ran the following regression with two dummy variables:

Client's happiness

$$= b_0 + b_1 * I_1 \text{ (client's estimate − feedback)}$$

$$+ b_2 * I_2 \text{ (feedback − client's estimate)},$$

where $I_1 = 1$ when (client's estimate − feedback) > 0 and $I_1 = 0$ otherwise, and $I_2 = 1$ when (feedback − client's estimate) > 0 and $I_2 = 0$ otherwise. The regression results show that the client's reported happiness dropped when the feedback was below expectations ($b_1 = -0.96$; $t(37) = -2.33$; $p < 0.05$) and remained unchanged when the feedback was above expectations ($b_2 = 0.28$; $t(37) = 0.73$; $p = 0.47$). Also, most agents tried to make their feedback believable by staying within a threshold of feedback inflation (over 90% of the feedback corresponded to or was less than four units higher than the actual performance).

Figure 3 Scatter Plot of Agent's Feedback Against Client's Performance Estimate in the Prop Condition in Study 5
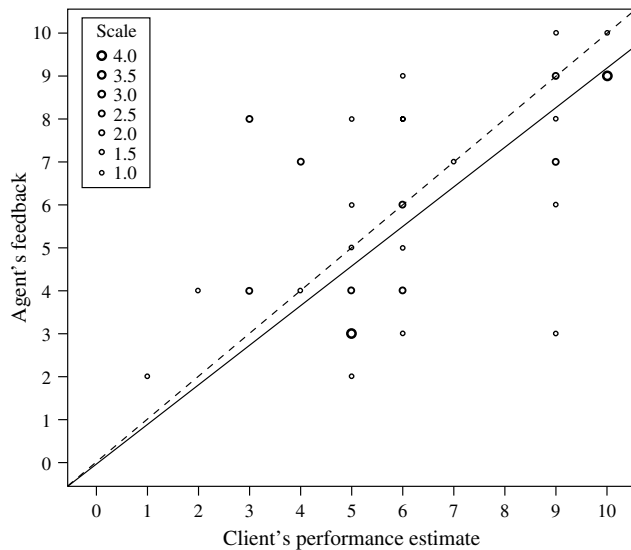
Notes. The solid line represents the best fitted line: feedback = 0.919 ∗ performance estimate. The dotted line represents the 45° line: feedback = performance estimate.

hypothesis by examining whether the opportunistic agent will stop inflating his feedback when his earnings are proportional to the client's ex post happiness.

6.2. Study 6: Testing the Agent's Opportunism
This study included the same two experimental conditions as Study 1 (the prop condition and the truth condition), except in both conditions the client was asked to report her happiness a second time, after she had found out her actual quiz performance, and the agent in the prop condition was paid $0.50 for each increased point in the client's ex post happiness. We refer to this new condition as the "prop-ex post" condition. The payment scheme for the truth condition was identical to that in Study 1; that is, the agent received $3 for truthful reporting and received nothing ($0) for untruthful reporting. Ninety-six undergraduates participated in this experiment. They were randomly assigned to either the prop-ex post condition or the truth condition.

6.2.1. Results. As shown in Table 2, the math quiz results were not statistically different between the two conditions ($M_{prop\text{-}ex\ post} = 5.66$ versus $M_{truth} = 5.03$; $F(1, 94) = 1.67$; $p = 0.20$). All the agents in the truth condition reported honestly. In the prop-ex post condition, a large majority of the agents (67.2%) reported honestly, and only a small fraction (9.4%) inflated their feedback. Interestingly, a significant minority (23.4%) deflated (i.e., underreported). The mean feedback value did not differ between the two conditions ($M_{prop\text{-}ex\ post} = 5.39$ versus $M_{truth} = 5.03$; $F(1, 94) < 1$). These results suggest that the agent was indeed opportunistic and chose not to inflate his feedback

when his earnings were proportional to the client's ex post happiness.

Client's Reported Happiness. The client's reported happiness before she found out her true performance was not statistically different in the two conditions ($M_{prop\text{-}ex\ post} = 5.09$ versus $M_{truth} = 4.34$; $F(1, 94) = 1.22$; $p = 0.27$). Interestingly, the client's ex post happiness rating was higher in the prop-ex post condition than in the truth condition ($M_{prop\text{-}ex\ post} = 5.73$ versus $M_{truth} = 4.22$; $F(1, 94) = 4.83$; $p < 0.05$).

Recall that 23.4% of the agents deflated their feedback. To determine whether feedback deflation increased the client's ex post happiness rating and hence the agent's earnings, we regressed the ex post happiness rating against feedback deflation. There was no evidence that feedback deflation affected the ex post happiness rating (beta = 0.31; $t(62) = 1.04$; $p = 0.30$). In fact, when the client's actual math quiz performance was added to the regression as another independent variable, the result clearly showed that the actual performance (beta = 0.56; $t(61) = 3.15$; $p < 0.01$) and not feedback deflation (beta = 0.28; $t(61) = 1.03$; $p = 0.31$) determined the ex post happiness. Thus, although some agents considered deflation to be a good strategy, the clients did not seem to respond to this preemptive tactic, and instead reacted mainly to their actual performance.

6.2.2. Discussion. The results of Studies 5 and 6 suggest that feedback inflation is driven by the agent's opportunistic motive; that is, the agent inflates his feedback only if feedback inflation is found to be financially rewarding. The agent receives a higher payoff when he inflates his feedback because the client mistakenly believes the (inflated) feedback and hence reports a higher level of happiness. This inflated feedback appears credible because the agent's feedback inflation closely matches the degree of the client's overestimation of her own performance.

We now move on to three additional experiments that examine the generalizability and robustness of our main findings. In Studies 7A and 7B, we examine whether our main findings hold if we vary two important contextual features of the feedback-giving game. In Study 8, we examine whether the agent continues to inflate his feedback if feedback inflation could have a negative effect on the client's monetary payoff.

7. Do Contexts Matter?
In Studies 7A and 7B, we changed two important contextual features of the feedback-giving game to check the robustness of the main empirical regularities. In Study 7A, we ran the feedback-giving game in an

abstract setting (i.e., in the tradition of experimental economics) by removing the referral to "happiness" from the experimental instructions entirely.[19] In Study 7B, we used a binary scale to present the client's performance and the agent's feedback (good versus bad), such that feedback inflation requires inverting the truth.

7.1. Study 7A: Feedback-Giving Game in Abstract Terms

Our feedback-giving game is designed to help us understand how people strategically manage others' reported happiness. Thus, one fundamental principle that we must capture in the game is the agent's strategic motivation to manage the client's reported happiness through feedback giving. We made this motivation transparent to both the agent and the client by asking the client to overtly report her happiness and by tying the agent's payoff to the client's reported happiness in a precise way. In Study 7A, we tested whether the agent would still inflate his feedback if the client did not report her happiness, but instead determined the agent's payoff by choosing a number on a scale from 0 to 10, knowing that each scale point would earn the agent $0.50. Thus, it was left to the agent to interpret whether or not his feedback would affect the client's determination of his payoff and, if so, how.[20] We refer to this experimental condition as the "abstract condition" because the game was conducted in an abstract setting in which the happiness label was removed from the experimental instructions and response scales. Eighty-four subjects participated in this experiment and were randomly assigned to either one of the two experimental conditions: the prop condition (as in the previous studies) and the abstract condition. Note that the two conditions were identical except that there was no mention of happiness whatsoever in the abstract condition.

7.1.1. Results. Our findings showed no difference between the two conditions in terms of math quiz performance ($M_{\text{prop}} = 4.68$; $M_{\text{abstract}} = 4.30$; $F < 1$).

Of the 40 agents in the prop condition, 42.5% inflated their feedback, and the other 57.5% reported honestly. Of the 44 agents in the abstract condition, 29.5% inflated, 68.2% reported honestly, and one agent underreported. There was no significant difference in the proportions of feedback inflation between the conditions (42.5% versus 29.5%; $z = 1.24$; $p = 0.21$), although the average degree of feedback inflation was marginally higher in the prop condition than in the abstract condition ($M_{\text{prop}} = 1.75$ versus $M_{\text{abstract}} = 0.93$; $F(1, 82) = 2.76$; $p = 0.10$). Most importantly, the average feedback was statistically higher than the actual quiz performance in both the prop (6.43 versus 4.68; $t(39) = 4.45$; $p < 0.001$) and abstract conditions (5.23 versus 4.30; $t(43) = 3.07$; $p < 0.001$), suggesting that the agent inflated his feedback in both settings. Thus, the agent's management of payoff through feedback inflation is likely to be a generic form of thinking that transcends the contextual setting in which the happiness label was used.

The agent's payoff was higher in the abstract condition ($7.25/2 = \$3.6$) than in the prop condition ($5.43/2 = \$2.7$; $F(1, 82) = 8.22$; $p < 0.01$); that is, the client became more generous toward the agent (with the experimenter's money) when her decision on the agent's payoff was reported on a pay scale without happiness level. (Recall that a similar pattern was observed in the "decoupled" condition in Study 2.) Put differently, combining the client's decision on the agent's payoff and on her reporting of happiness can limit the client from being generous with the agent.

7.2. Study 7B: Giving Binary Feedback

Past research shows that opportunistic acts of self-reporting, such as overreporting one's performance in a task or overclaiming the cost of a repair bill to receive a higher reimbursement, depend on the ease of morally justifying the misreporting (Mazar et al. 2008, Schweitzer and Hsee 2002). An opportunistic act of report inflation by a few units out of a wide range of possible outcomes can be easily justified as "stretching the truth" and is therefore not considered as dishonesty (Schweitzer and Hsee 2002). Similarly, the agents in our feedback-giving game may consider it justifiable to inflate their feedback because they regard a few units of inflation as truth-stretching rather than outright dishonesty. In Study 7B, we tested whether the agent continued to inflate his feedback in situations where the performance rating was binary (i.e., "bad" and "good") such that feedback inflation (from "bad" to "good") required inverting the truth.

One hundred sixteen participants were randomly assigned to one of the two experimental conditions: (1) the prop condition and (2) the truth condition (the same set of conditions as Study 1). The experimental

---

[19] Experimental economists prefer to investigate behavior in abstract settings to avoid additional "induced value" associated with any context labeling. We intentionally use the happiness label in the experimental instructions because we are interested in people's strategic management of others' reported happiness and believe this explicit labeling helps to increase the external validity of our empirical regularities.

[20] We thank a reviewer for pushing us to conduct this additional study. Note that our feedback-giving game is a signaling game. In this game, the agent receives some private information and must send a message to the client. The client then responds by choosing an action that can influence the agent's monetary payoff. One can potentially derive a Bayesian Nash equilibrium prediction in the resulting signaling game. However, this analysis is beyond the scope of this paper.

procedure was similar to that in Study 1, except the client's math quiz performance was converted into the binary categories of "good" or "bad" by a Web-based software program. This binary categorization applied to both the information the agent received about the client's performance ("Your paired member's performance is ——") and the feedback he conveyed to the client ("Your performance is ——"). It was common knowledge that performance would be coded as "bad" when the number of correct answers ranged from 0 to 5 and as "good" when the number of correct answers ranged from 6 to 10. This information was included in the experimental instructions and shown on the screen where the agent read about the client's performance. As in Study 1, the agent was paid either based on the client's reported happiness (prop condition) or based on whether he reported truthfully (truth condition). Like before, the client was paid $1 for each math question she answered correctly.

The experimental results are shown in Table 2. Of the 116 clients, 58 had quiz performance that fell into the "good" category, and no inflation could possibly take place in the feedback giving. Therefore, of particular interest here was whether the remaining 58 agents, whose clients had a quiz performance that fell into the "bad" category, would give the inflated feedback of "good." As expected, none of these agents in the truth condition inflated their feedback. However, more than half of the agents in the prop condition (18 of 35) did so by inverting the truth. The feedback inflation in the prop condition resulted in the clients in the prop condition reporting a marginally higher level of happiness than those in the truth condition ($M_{prop} = 7.22$ versus $M_{truth} = 6.14$; $F(1, 114) = 3.17$; $p = 0.08$).

8. Study 8: Giving Consequential Feedback

In Studies 1 to 7, the agent's feedback bore no monetary consequence to the client, because the client was paid based on her performance in a task that had already been completed before the feedback was received. In Study 8, we examined feedback giving in a setting where the agent's feedback was consequential in that it could influence the client's decision on how she wished to be compensated (i.e., a fixed fee independent of her task performance versus a fee proportional to her task performance) in an additional but identical task performed right after receiving the agent's feedback. Specifically, the client had to perform two math quizzes that were identical in terms of format and level of difficulty (i.e., to finish 10 SAT-type math questions within 15 minutes). The client was compensated only for the second

math quiz and had to choose between being compensated (a) $1 for each correctly answered question or (b) a fixed fee of $5 for simply undertaking the second quiz. The choice of payment scheme was made after the client had completed and received feedback from the agent about her performance in the first math quiz but before the second quiz began. Since the client's performance in the first quiz is highly predictive of her performance in the second quiz, an honest feedback can really help her choose the right payment scheme. However, inflated feedback may mislead an overconfident client to choose the performance-based payment scheme, which may result in a lower payoff. Most importantly, the agent was made keenly aware of the potentially negative consequence of his lie, because he anticipated the client would believe his feedback (as long as his feedback strategy was to lie only to the extent that the feedback remained believable, as shown in the previous studies).

Eighty participants participated in this study. There was only one experimental condition, in which the clients were asked to choose how they wished to be paid in the second quiz. All of the agents were paid based on the level of happiness the clients reported after receiving feedback on the first quiz. Each scale point earned the agent $0.50.

8.1. Results
We examined whether the agent inflated his feedback despite knowing the potentially negative consequence his feedback could have on the client's monetary payoff. Of the 80 agents, 61% reported honestly, 38% inflated, and 1% (one agent) deflated. The clients answered an average of 5.51 questions correctly in the first quiz, and the feedback they received was inflated by 1.18 units (significantly different from 0; $t(79) = 5.53$; $p < 0.01$). We were particularly interested in cases where the clients answered less than five questions correctly. For this group of clients, an inflated feedback that went beyond 5 would have been detrimental because it would have pointed the clients to the performance-based payment scheme instead of the more profitable fixed-fee payment scheme. Our data show that 41 clients answered fewer than five questions correctly in the first quiz, of which 15 (37%) received feedback that had been inflated beyond 5. This result clearly shows that the agent continued to inflate his feedback in this two-stage game, despite its potentially detrimental effect on the client's cash earnings. (Other findings on the client's reported happiness are reported in Table 2. Because these findings are consistent with those reported in the other studies, they are not discussed here.)[21]

---

[21] To check whether the client's choice of compensation scheme was influenced by the agent's feedback before the client knew about

9. General Discussion

The series of eight tightly linked studies reported in this paper investigates a prevalent client–agent interaction using a new feedback-giving game. In this game, the client performs a task and her performance is disclosed to the agent. The agent must formulate feedback to inform the client of her performance. Upon receiving feedback and before knowing her true performance, the client reports her level of happiness, which in turn determines the agent's monetary payoff: the higher the client's reported happiness, the higher the agent's payoff. We investigate whether feedback inflation occurs in this game and, if so, whether it pays, and if it pays, why.

We systematically vary the relationship between the agent's monetary payoff and the client's reported happiness. We show that the agent inflates his feedback when his cash earnings increase linearly with the client's reported happiness. The agent receives a higher payoff when he inflates his feedback because the client mistakenly believes the (inflated) feedback is genuine and hence reports a higher level of happiness. This inflated feedback appears credible because the agent's inflation closely matches the degree of the client's overestimation of her own performance.

We rule out the altruism hypothesis with respect to both the client and the agent. The agent does not tell a white lie by inflating his message to make the client happier. The client does not report a higher level of happiness to help the agent make more money. Also, we reject the conditional altruism hypothesis that the client is altruistic only if she is ahead of the agent in terms of payoff. Finally, we observe that the agent stops inflating his feedback when we make his cash earnings proportional to the client's level of happiness reported *after* she has found out her actual performance. This finding confirms that the agent is indeed opportunistic and will inflate his feedback only if he can benefit by doing so (i.e., if happiness reporting takes place before the client finds out her true performance). The main empirical regularities are replicated in three new contexts: (1) in an abstract context where the happiness label is removed, (2) in a context where the client's performance and the agent's feedback are both binary (i.e., good versus bad), and (3) in a context where the agent's feedback can potentially

influence the client's payoff.[22] Thus, the main empirical regularities appear quite robust.

9.1. Real-Life Feedback Giving and Receiving

Our research findings were obtained under a controlled experiment in which the client's performance was objectively defined and measured, and the true performance was revealed to the client within the time frame of the experimental session. The task design makes the client's incentive mechanism transparent and her task performance consequential. Interestingly, the same design also sets a stringent condition for the agent to distort his feedback. Both the existence of a well-defined objective truth and the short elapsed time for this truth to be revealed make it morally difficult to distort the client's performance and emotionally hard to ultimately disappoint the client. In this regard, our experimental setting provides a conservative test of feedback inflation. Given that the agent inflates his feedback under this setting, one would expect him to be even more likely to do so in a real-world setting where there is no well-defined objective truth or where it takes a long time for the truth to materialize, if ever.

When truth is not well defined, the agent and the client are both left in a gray area to determine the truthfulness of any feedback. Specifically, the client may find it difficult to verify whether the agent is acting honestly or dishonestly. For example, when a salesperson is asked to evaluate the stylistic fit of a product by a potential customer, the customer can never tell whether the salesperson is being truthful or not because there is no objective truth in such evaluation. Similarly, there is no way for a boss to tell what her subordinate really thinks when the latter is asked to provide feedback on her personal traits. In both cases, the lack of objective truth reduces the negative consequences of feedback inflation and makes the agent more likely to take advantage of the situation by inflating his feedback.[23]

When the truth takes a long time to materialize, ex ante happiness (i.e., before the truth is known) is more salient and important than ex post happiness in client–agent interactions. For example, a junior colleague is likely to inflate his feedback when assessing a senior colleague's research paper because the true quality of the paper can only be imperfectly determined by a lengthy publication review process. The senior colleague may find out the true quality eventually, but the junior colleague will, meanwhile, enjoy

---

her true performance in the first quiz, we ran a logistic regression using the client's choice of compensation scheme as the dependent variable and the agent's feedback as the independent variable. The result showed that the probability of choosing the performance-based payment scheme (i.e., $1 for each correct answer) increased with the favorability of the feedback ($\beta = 0.31$; $\chi^2(1) = 7.27$; $p < 0.01$). Thus, it is evident that the client used the agent's feedback as a way to decide how she wished to be compensated in the second math quiz.

---

[22] In Study 8, we also show that feedback affects the client's subsequent behavior (choice of payment option). For more in-depth discussion on how feedback may result in learning and behavioral change, see Bandura (1971).

[23] For related works on agent–client interactions where the client is unlikely to find out the truth, see Cain et al. (2011).

the support of the senior colleague in a forthcoming promotion. Here, the ex ante happiness of the senior colleague drives the strategic behavior on the part of the junior colleague.

The above examples also illustrate the practical relevance of tying the agent's payoff to the client's ex ante happiness. In these examples, the ex post happiness is either nonexistent (when there is no well-defined objective truth) or is dominated by the ex ante happiness in determining the agent's payoff (when there is a long elapsed time for the truth to be revealed). Thus, the explicit link between the agent's payoff and the client's ex ante happiness is highly representative of the manner in which agents are compensated in real-life situations.[24]

9.2. Limitations and Future Research
Our paper has several limitations and allows room for possible extensions. First, our studies assume a one-shot client–agent interaction. If the feedback communication stage is repeated several times between the same agent and client, and the actual performance is disclosed to the client each time, then one would expect the client to learn about the agent's tendency to inflate his feedback and hence discount it accordingly. As a consequence, feedback inflation might not pay in this repeated setting where there is room for learning and reputation building. Future research can examine this issue explicitly. Second, the agent in our feedback-giving game has perfect knowledge of the client's performance. In some client–agent interactions, the agent might have imperfect knowledge of the client's actual performance. Thus, it would be interesting to examine whether this imperfect knowledge on the part of the agent increases or decreases the extent of his feedback inflation. Our experimental paradigm can easily be extended to shed light on this practical setting. For example, we can inform the agent on how well the client does in only a subset of the math quiz problems (e.g., the number of correct answers in the first 5 of the 10 questions) and ask him to provide feedback on the client's overall performance. Third, our agent is only allowed to provide precise feedback (*x* of the 10 questions are answered correctly). In practice, feedback can be imprecise and ordinal (e.g., poor, average, good). Therefore, it would be worthwhile to study how the degree of inflation changes in providing ordinal feedback. Fourth, it would be useful to generalize the feedback-giving game to a setting where there are multiple agents who provide feedback to the same

client. Here, an individual agent who provides feedback to a client must consider how his feedback interacts with that of others and how the client will report her happiness toward the agent when his feedback is either more positive or more negative than that of others.

Appendix. Experimental Instructions for Study 1 (Prop Condition)
This is an experiment on decision making. The instructions are simple, and if you follow them carefully, there is a chance that you could earn a considerable amount of money, which will be paid to you in cash before you leave today.

The experiment will consist of two parts. In the first part, all subjects will undertake a *math quiz*. Subjects will be compensated based on their performance on the math quiz. In the second part, subjects will be paired in groups of two *twice*. The matching protocol ensures that each subject will be paired with a different subject in each of the two pairings (i.e., you will not be matched with the same subject twice). Subjects will communicate to their paired members about the latter's performance in the math quiz in one pairing and will receive feedback from their paired members about their own performance in the math quiz in another pairing; that is, subjects will provide feedback about others' performance in one pairing and receive feedback from others about their own performance in another pairing. The order of these two communication episodes (i.e., sending feedback and receiving feedback) is determined by a computer-simulated coin toss and depends solely on chance. *Subjects will also be compensated when they provide feedback, and the details of the compensation will be discussed in the following pages.*

Any communications between subjects in each pair must be made through a Web-based computer program. It is important that you do not look at each other, and that you do not talk, laugh, or exclaim aloud during the experiment. You will be warned if you violate this rule the first time. If you violate this rule twice, you will be asked to leave, and you will not be paid; that is, your earnings will be $0.

---

[24] This explicit link is common knowledge in all of our experiments. In situations where the agent's incentive is not common knowledge (i.e., when he knows how he is paid, but the client does not know the agent's incentive), the client is even more likely to believe in the agent's feedback.

Experimental Procedure

Again, the experiment consists of two tasks: (1) a *math quiz* and (2) a *social communication task*. Your total cash payment will be the *sum of the money you earn in each of these tasks*.

Math quiz. Every subject will undertake a math quiz in the first part of the experiment. You will answer 10 standard SAT math questions (SAT stands for "Scholastic Aptitude Test," a standardized test that is often used to assess a student's readiness for college. In other words, all college students should be reasonably capable of solving these questions). Your answers will be graded by the computer program. *You will receive* $1 *as a reward for each correct answer*. In other words, you can earn up to $10 based on the math quiz, and the money will be paid to you in cash before you leave the lab today.

Social communication task. As mentioned, your answers in the math quiz will be graded by the computer program. However, you will not find out your performance directly from the program. Your performance report will be sent to a subject who is randomly paired with you, and this subject will give you feedback on your performance. Likewise, you will receive the performance report of a randomly paired subject and you will give feedback to this paired member. The order of these two communication episodes is determined by a computer-simulated coin toss and depends solely on chance. The exact stpdf involved in this task and the payoff scheme are outlined next.

You give feedback to your paired member. After the system has finished grading all the tests, you will receive a performance report from the program, indicating your paired member's performance in the math quiz. The report should look like the one below:

Question	Answer
1	Correct
2	Incorrect
3	Incorrect
4	Correct
5	Correct
6	Correct
7	Correct
8	Correct
9	Incorrect
10	Incorrect

Your job is to give feedback to your paired member ("You have correctly answered —— questions").

Your paired member will receive this feedback and indicate how happy he or she is on the following scale:

| 0	1		2	3	4	5	6	7	8	9		10
Not happy at all												Very happy

You receive feedback from a different paired member. The procedure is identical to the above except that each subject will be *paired up with a different paired member* and the *roles will be switched* (i.e., you will receive feedback from your paired member). More specifically,

• Your paired member will receive your performance from the computer program. His or her job is to provide feedback to you ("You have correctly answered —— questions").

• You will then read this feedback and indicate how happy you are on the scale (0–10) mentioned earlier.

All subjects retrieve their own performance report. All subjects will be provided with their own actual performance report from the computer program.

The following is a flow chart summarizing the procedure of the second part of the experiment:
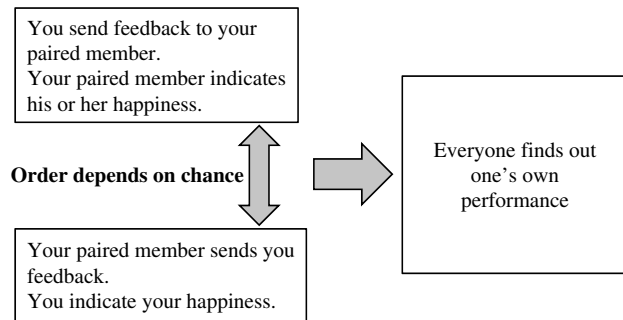
Payoff. As indicated above, you will be compensated for both parts of the experiment. Your payment for the math quiz will be based on your actual performance in the quiz, and you will receive $1 for each correct answer in the quiz.

In the social communication task, you will be paid by providing feedback, and your payoff will *depend on the paired subject's happiness right after reading your feedback*. As mentioned before, your paired subject will indicate his or her happiness on a scale from 0 (not at all happy) to 10 (very happy). *Each scale point will yield a payoff of* $0.50. For example, if your paired member indicates a happiness rating of "10," you will receive $5 (10 × 0.5) for the social communication task. If your paired member indicates "0," you will receive nothing for the social communication task. Similarly, a happiness rating of 5 will give you a payoff of $2.50. In other words, the higher your paired member's reported happiness after reading your feedback, the higher the payment will be.

Your total payoff for today's experiment is the total amount of money you earn from the math quiz and the social communication task. You will receive the payment in cash when you leave the experiment.

References

Andrade E, Ho TH (2007) How is the boss's mood today? I want a raise. *Psych. Sci.* 18(8):668–671.

Andrade E, Ho TH (2009) Gaming emotions in social interactions. *J. Consumer Res.* 36(4):539–552.

Andreoni J (1990) Impure altruism and donations to public goods: A theory of warm-glow giving. *Econom. J.* 100(401):464–477.

Bandura A (1971) *Social Learning Theory* (General Learning Press, New York).

Bolton G, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *Amer. Econom. Rev.* 90(1):166–193.

Cain DM, Loewenstein G, Moore DA (2011) When sunlight fails to disinfect: Understanding the perverse effects of disclosing conflicts of interest. *J. Consumer Res.* 37(5):836–857.

Camerer CF (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press, Princeton, NJ).

Eil D, Rao JM (2011) The good news–bad news effect: Asymmetric processing of objective information about yourself. *Amer. Econom. J.: Microeconomics* 3(2):114–138.

Erat S, Gneezy U (2012) White lies. *Management Sci.* 58(4): 723–733.

Forsythe R, Horowitz JL, Savin NE, Sefton M (1994) Fairness in simple bargaining experiments. *Games Econom. Behav.* 6(3): 347–369.

Gneezy U (2005) Deception: The role of consequences. *Amer. Econom. Rev.* 95(1):384–394.

Gordon RA (1996) Impact of ingratiation on judgments and evaluations: A meta-analytic investigation. *J. Personality Soc. Psych.* 71(1):54–70.

Kruger J, Burrus J (2004) Egocentrism and focalism in unrealistic optimism (and pessimism). *J. Experiment. Soc. Psych.* 40(3): 332–340.

Mazar N, Amir O, Ariely D (2008) The dishonesty of honest people: A theory of self-concept maintenance. *J. Marketing Res.* 45(6):633–644.

Mobius MM, Niederle M, Niehaus P, Rosenblat TS (2011) Managing self-confidence: Theory and experimental evidence. NBER Working Paper 17014, National Bureau of Economic Research, Cambridge, MA.

Moore DA, Healy PJ (2008) The trouble with overconfidence. *Psych. Rev.* 115(2):502–517.

Moore DA, Small DA (2007) Error and bias in comparative judgment: On being both better and worse than we think we are. *J. Personality Soc. Psych.* 92(6):972–989.

Moore DA, Kurtzberg TR, Fox CR, Bazerman MH (1999) Positive illusions and forecasting errors in mutual fund investment decisions. *Organ. Behav. Human Decision Processes* 79(2):95–114.

Schweitzer ME, Hsee CK (2002) Stretching the truth: Elastic justification and motivated communication of uncertain information. *J. Risk Uncertainty* 25(2):185–201.

Sedikides C (1993) Assessment, enhance, and verification determinants of the self-evaluation process. *J. Personality Soc. Psych.* 65(2):317–338.

Sugden R (1982) On the economics of philanthropy. *Econom. J.* 92(366):341–350.

Taylor SE, Brown JD (1988) Illusion and well-being: A social psychological perspective on mental health. *Psych. Bull.* 103(2): 193–210.